



MITE PROTEIN

Technical field

The present invention relates to a novel recombinant protein useful in the diagnosis of mite infections in mammals, such as dogs. The invention also relates to a nucleic acid encoding the protein according to the invention as well as to various advantageous uses thereof.

Background

Scabies or sarcoptic mange, which is caused by infection with the parasitic mite *Sarcoptes scabiei*, is a widespread, highly contagious disease. The parasite has been found in well over 40 different mammals, including man. During the infection, female mites burrow in the skin to feed and their activities cause an intense irritation that leads to itching and scratching, which may aggravate the condition. Sensitization of the host to the mites and their products probably plays an important role in the pathogenesis of the disease. In its extreme form, scabies can develop into a severe hyperkeratotic form, where several thousands of mites can be found in the lesions. This form of crusted scabbies is also frequently observed among immune compromised individuals, *i.e.* HIV patients.

The standard method for diagnosing scabies include microscopic detection of the mites and their eggs and feces in skin scrapings. This method is relatively time-consuming and, in many cases, the results of microscopic examination can be negative owing to the low number of parasites present in each sample. The use of enzyme-linked immunosorbent assays (ELISA) for the detection of antibodies to *S. scabiei* has been reported for several different animals including humans (Bornstein & Zakrisson (1993) *Vet. Dermatol.* 4:107; Arlian *et al.* (1994) *Exp. Parasitol.* 78:37; Normaznah *et al.* (1996) *Southeast Asian J. Trop. Med. Public Health* 27:53; and Hollanders *et al.* (1997) *Vet. Parasitol.* 69:117). The ELISA technology offers several distinct advantages compared to skin scrapings: A relative ease of sampling, the potential to screen large numbers of samples and the possibility to standardize

reagents and methods. The only major drawback is the limited amount of antigenic material available due to the lack of an *in vitro* propagation system for *S. scabiei*.

Arlian and colleagues (J. Med. Entomol. (1988) 25:52) have developed an *in vivo* propagation system after establishing *S. scabiei* var. *canis* on rabbits. However, the parasite burden is still relatively low and the method cannot be applied for the production of antigens for large scale screening projects. Mites from naturally infected red foxes have been successfully isolated in large numbers and used in ELISAs both for dogs and pigs. The difficulties involved in using wild foxes as a source for antigen supply are however evident.

This lack of material has not only limited large scale screening efforts and control programs among farm animals, but it has also limited the possibilities to study other aspects of scabies, for instance pathogenesis.

EP 0 473 111 provides a recombinant mite allergen which is effective as a therapeutic agent and diagnostic reagent for mite allergic diseases. However, the mite allergen of this reference are derived from a house dust mite, more specifically from *Dermatophagoides farinae*. Thus, firstly, the said house dust mite is found free-living in a house-hold environment and does not live as a parasite on a host. Thus, disease thereof is not the result of an infection but the inhalation of mites or mite body fragments or faeces. The inhalation of house dust mites may cause allergic reactions in susceptible individuals resulting in respiratory distress, such as asthma and rhinitis. Accordingly, EP 0 473 111 relates to molecules useful in the diagnosis and prevention of such conditions.

Summary of the invention

According to the present invention, the problems defined above are solved by providing a novel major mite antigen, which has now been isolated and sequenced for the first time. Thus, the invention relates an isolated antigenic protein as well as

to the nucleic acid said novel protein, as defined in the appended claims. Further, the invention also relates to various advantageous uses of the novel protein and functional fragments thereof, *e.g.* in immunological testing, such as in ELISA methods.

5

Brief description of the drawings

Figure 1 shows the extent of the recombinant protein according to the invention in relation to the native protein.

10

Figure 2 illustrates the recombinant plasmid pPU17 used for the expression of the mite protein according to the invention.

Figure 3 is an overview of the cloning strategy of the 5' end of MSA.

Figure 4 shows the results of Western blot analysis of the mite recombinant protein according to the invention.

15

Definitions

20

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

25

The term "nucleic acid" refers to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, encompasses known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides.

30

The phrase "hybridizing specifically to", refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular DNA or RNA. The term "stringent conditions" refers to conditions under which a probe will hybridize to its target subsequence, but to no other sequences. Stringent

conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point T_m for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. (As the target sequences are generally present in excess, at T_m , 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ^{32}P , fluorescent dyes, electron-dense reagents, enzymes (*e.g.*, as commonly used in a ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available.

As used herein a "nucleic acid probe" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.* A, G, C, U or T) or modified bases (7-deazaguanosine, inosine, *etc.*) In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization.

A "labeled nucleic acid probe" is a nucleic acid probe that is bound, either covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a

label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

5 The term "target nucleic acid" refers to a nucleic acid (often derived from a biological sample), to which a nucleic acid probe is designed to specifically hybridize. It is either the presence or absence of the target nucleic acid that is to be detected, or the amount of the target nucleic acid that is to be quantified. The target nucleic acid has a sequence that is complementary to the nucleic acid sequence of the corresponding probe directed to the target. The term target nucleic acid may
10 refer to the specific subsequence of a larger nucleic acid to which the probe is directed or to the overall sequence (*e.g.*, gene or mRNA) whose expression level it is desired to detect. The difference in usage will be apparent from context.

15 "Subsequence" refers to a sequence of nucleic acids or amino acids that comprise a part of a longer sequence of nucleic acids or amino acids (*e.g.*, polypeptide), respectively.

20 "Combinatorial library" means a library of molecules containing a large number, typically between 10^3 and 10^6 , of different-sequence oligomers, typically characterized in having different sequences or subunits, or a combination of different sequences of side chains and linkages, or different-substituent compounds in a small-compound library.

25 "Different-sequence oligomer compounds" are oligomers, such as oligonucleotides, oligonucleotide analogs, oligopeptides, oligopeptide analogs, oligosaccharides, or lipopeptides with different permutations of lipid and/or sequences in the peptide moieties, glycopeptides with different sequence permutations in the saccharide and/or peptide moieties, non-biological oligomers with different-sequence permutations, or different-substituent compounds in a small-compound library.

The term "recombinant" when used with reference to a cell, or nucleic acid, or vector, indicates that the cell, or nucleic acid, or vector, has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. Further, the expression "recombinant" also relates to a cell, wherein further regulatory elements have been included in order to initiate or enhance expression of an otherwise silent endogenous gene, or wherein a manipulation of the regulatory elements have been performed for the same purpose. (For an example of such a gene activation technique, see *e.g. Genetic Engineering News*, April 15, 1994.)

The term "identical" in the context of two nucleic acid or polypeptide sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2: 482, by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI) or by inspection.

An additional algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). (See Henikoff and Henikoff (1992) *Proc. Natl. Acad. Sci. USA* 89: 10915-10919; and Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90: 5873-5787; for further information in this context.)

The term "substantial identity" or "substantial similarity" in the context of a polypeptide indicates that a polypeptides comprises a sequence with at least 70% sequence identity to a reference sequence, or preferably 80%, or more preferably 85% sequence identity to the reference sequence, or most preferably 90% identity over a comparison window of about 10-20 amino acid residues. An indication that two polypeptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a polypeptide is substantially identical to a second polypeptide, for example, where the two peptides differ only by a conservative substitution.

An indication that two nucleic acid sequences are substantially identical is that the polypeptide which the first nucleic acid encodes is immunologically cross reactive with the polypeptide encoded by the second nucleic acid. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions.

"Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accomodated by reducing the stringency of the hybridization media to achieve the desired detection of the target polynucleotide sequence.

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. (In this context,

see *Fundamental Immunology*, Third Edition, W.E. Paul, ed., Raven Press, N.Y. 1993).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, *etc.*; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

The term "immunoassay" is an assay that utilizes an antibody to specifically bind an analyte. The immunoassay is characterized by the use of specific binding properties of a particular antibody to isolate, target, and/or quantify the analyte.

The terms "isolated" "purified" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany it as found in its native state.

The phrases "specifically binds to a protein" or "specifically immunoreactive with", when referring to an antibody refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind preferentially to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbour Publications, New York,

for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

A "conservative substitution", when describing a protein, refers to a change in the amino acid composition of the protein that does not substantially alter the protein's activity. Thus, "conservatively modified variations" of a particular amino acid sequence refers to amino acid substitutions of those amino acids that are not critical for protein activity or substitution of amino acids with other amino acids having similar properties (*e.g.*, acidic, basic, positively or negatively charged, polar or non-polar, *etc.*) such that the substitutions of even critical amino acids do not substantially alter activity. See *e.g.*, Creighton (1984) *Protein*, W.H. Freeman and Company. In addition, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids in an encoded sequence are also "conservatively modified variations".

A "gene product", as used herein, refers to a nucleic acid whose presence, absence, quantity, or nucleic acid sequence is indicative of a presence, absence, quantity, or nucleic acid composition of the gene. Gene products thus include, but are not limited to, an mRNA transcript, a cDNA reverse transcribed from an mRNA, and RNA transcribed from that cDNA, a DNA amplified from the cDNA, an RNA transcribed from the amplified DNA or subsequences of any of these nucleic acids. Polypeptides expressed by the gene or subsequences thereof are also gene products. The particular type of gene product will be evident from the context of the usage of the term.

Detailed description of the invention

In a first aspect, the present invention relates to an isolated mite-derived protein, which is an antigen derived from the mite *Sarcoptes scabiei*. More specifically, the protein according to the invention is comprised of parts or all of the sequence disclosed in SEQ ID NO. 2. More specifically, the invention relates to a protein comprised of at least about 83 amino acids of said sequence, preferably the ones

described in SEQ ID NO. 3, an analogue or a functional fragment thereof. In a specific embodiment, the protein according to the invention is comprised of at least about 100 amino acids, preferably at least about 200, such as at least about 400 amino acids of the sequence disclosed in SEQ ID NO. 2. In a specific embodiment, the protein according to the invention comprises about 400 of the last amino acids of SEQ ID NO 2, and most preferably, the sequence from about amino acid no 344 to amino acid no 770 of said sequence, which protein has been denoted Major Sarcoptes Antigen 1 (MSA1) of the present inventors. In another particular embodiment, the amino acid sequence of the present protein comprises a larger part of said sequence, such as at least about 400, preferably at least about 500, e.g. at least about 600, and most preferably at least about 700, such as about 770 amino acids of said sequence, in which case it is substantially identical with the sequence disclosed in SEQ ID NO. 2. In this context, it is to be understood that all derivatives, analogues and functional fragments and functional subsequences thereof of the present proteins also fall within the scope of the invention as defined by the claims.

Thus, the above mentioned EP 0 473 111 relates to the isolation of an organism, house dust mite, which is naturally present in different surroundings than *Sarcoptes scabiei*, which is a parasite in the skin of animals, such as wild foxes, swine etc., and man. As a consequence, the pharmaceutical and diagnostic applications enabled by the EP 0 473 111 antigen are suited for different conditions and other subjects than the antigen according to the present invention, which mainly relates to the treatment and/or diagnosis of farm animals, such as swine, or dogs. Accordingly, even though there exist a certain similarity between the protein derived from the house dust mites according to EP 0 473 111 and the *Sarcoptes scabiei* protein isolated and sequenced according to the present invention, the differences in environment and conditions to be prevented and/or diagnosed are of such importance that 0 473 111 must be regarded as relating to a different field from the present invention.

In an especially advantageous embodiment, the protein according to the invention is a recombinant protein, which in addition to a functional part or all of the sequence disclosed in SEQ ID NO. 2 also comprises a tag, such as a conventional further amino acid sequence, which confers properties that facilitates purification, downstream analysis, such as Western blot, reversible immobilization, immunoprecipitation, immunofluorescence analysis *etc.* Said tag may e.g. be the peptide His6. In a particular embodiment, the present invention is a fusion protein, wherein the present protein or a functional subsequence thereof is fused with another protein, such as β -galactosidase, glutathione-S-transferase, protein A *etc.* In the context of fusion proteins, see e.g. Smith and Johnson (1988) *Gene* 67:31; Hopp *et al.* (1988) *Biotechnology* 6:1204; La Vallie *et al.* (1993) *Biotechnology* 11:187.

In a further embodiment, the present protein, or a functional subsequence thereof, is used to produce peptidomimetics, *i.e.* molecules that mimics the biological activity of the peptide but no longer are completely peptidic in nature. Thus, peptidomimetics according to the invention are produced in order to provide molecules that are more advantageous than the peptides *per se* as regards e.g. size, bioavailability, duration of action, stability, storage, immunoreactivity *etc.* In this context, see e.g. Dean (1994), *BioEssays*, 683-687, Cohen and Shatzmiller (1993), *J. Mol. Graph.*, 11:166-173.

Further, the antigen according to the invention may be used in various other methods for design and/or identification of novel compounds, such as in combinatorial libraries. Such methods enable a modern drug laboratory to produce and screen millions of new chemical and/or biological compounds in a few weeks for a variety of uses. Of the large number of compounds produced, only the ones showing interesting biological activity are analyzed for further testing and experimentation. Methods for combinatorial libraries, see e.g. US patent nos. 5 753 187 and US 5 763 263, are well known to those of skill in this field and are easily

arranged based on the present application. A further aspect of the present invention is the use of the antigen according to the invention in methods of high-throughput screening for the identification of a novel compound. Such methods are amenable to automated, cost-effective high throughput screening and have immediate application in a broad range of programs for development of diagnostic and/or pharmaceutically active compounds. Further, the present invention also encompasses any compound obtainable by, that is, reached through, the use of any one of the herein disclosed methods.

Thus, more specifically, the present invention also relates to a method of screening for protein or peptide analogues that mimic at least a part of the structure of the protein according to the invention, which comprises the steps of

- (a) producing a multiplicity of analogue structures and
- (b) selecting an analogue structure, wherein the three-dimensional configuration and spatial arrangement of one or more biologically active regions, preferably antigenic regions, remain substantially preserved.

In a specific embodiment, analogues mimicking a protein having the amino acid sequence as disclosed in SEQ ID NO 3 are screened for. Further, the invention also encompasses an analogue identified as disclosed above. Analogues identified by the present method are advantageously used for the same purposes as the proteins according to the invention, said various uses being disclosed elsewhere in the present disclosure.

The present proteins, functional subsequences thereof etc. may be synthesized using standard chemical peptide synthesis techniques. Where the desired subsequences are relatively short (*e.g.*, when a particular antigenic determinant is desired), the molecule may be synthesized as a single contiguous polypeptide. Where larger molecules are desired, subsequences can be synthesized separately (in one or more units) and then fused by condensation of the amino terminus of one molecule with the carboxyl terminus of the other molecule thereby forming a peptide bond.

Solid phase synthesis in which the C-terminal amino acid of the sequence is attached to an insoluble support followed by sequential addition of the remaining amino acids in the sequence is the preferred method for the chemical synthesis of the polypeptides of this invention. Techniques for solid phase synthesis are described by Barany and Merrifield, *Solid-Phase Peptide Synthesis*; pp. 3-284 in *The Peptides: Analysis, Synthesis, Biology. Vol. 2: Special Methods in Peptide Synthesis, Part A.*, Merrifield, et al. *J. Am. Chem. Soc.*, 85: 2149-2156 (1963), and Stewart et al., *Solid Phase Peptide Synthesis, 2nd ed.* Pierce Chem. Co., Rockford, Ill. (1984).

In an alternative embodiment, the present proteins or subsequences thereof, are synthesized using recombinant DNA methodology. Generally this involves creating a DNA sequence that encodes the fusion protein, placing the DNA in an expression cassette under the control of a particular promoter, expressing the protein in a host, isolating the expressed protein and, if required, renaturing the protein.

DNA encoding the present proteins, or subsequences thereof, may be prepared by any suitable method, including, for example, cloning and restriction of appropriate sequences or direct chemical synthesis by methods such as the phosphotriester method of Narang et al. *Meth. Enzymol.* 68: 90-99 (1979); the phosphodiester method of Brown et al., *Meth. Enzymol.* 68: 109-151 (1979); the diethylphosphoramidite method of Beaucage et al., *Tetra. Lett.*, 22: 1859-1862 (1981); and the solid support method of U.S. Patent No. 4,458,066.

Chemical synthesis produces a single stranded oligonucleotide. This may be converted into double stranded DNA by hybridization with a complementary sequence, or by polymerization with a DNA polymerase using the single strand as a template. One of skill would recognize that while chemical synthesis of DNA is limited to sequences of about 100 bases, longer sequences may be obtained by the ligation of shorter sequences. Similarly, subsequences may be cloned and the

appropriate subsequences cleaved using appropriate restriction enzymes. The fragments may then be ligated to produce the desired DNA sequence.

In one embodiment, proteins of this invention may be cloned using DNA amplification methods such as polymerase chain reaction (PCR). Thus, for example, the nucleic acid sequence or subsequence is PCR amplified, using a sense primer containing one restriction site, *e.g.*, XhoI, and an antisense primer containing another restriction site, *e.g.*, BamHI. This will produce a nucleic acid encoding the desired sequence or subsequence and having terminal restriction sites. This nucleic acid can then be easily ligated into a vector containing a nucleic acid encoding the second molecule and having the appropriate corresponding restriction sites. Suitable PCR primers can be determined by one of skill in the art using the information provided in the appended sequence listing. Appropriate restriction sites can also be added to the nucleic acid encoding the protein or protein subsequence by site-directed mutagenesis. The plasmid containing the sequence or subsequence is cleaved with the appropriate restriction endonuclease and then ligated into the vector encoding the second molecule according to standard methods.

Accordingly, in a second aspect, the present invention relates to an isolated nucleic acid encoding the protein according to the present invention. More specifically, the present invention relates to a nucleic acid comprising parts or all of the base sequence disclosed in SEQ ID NO. 1, such as about 200-2000, *e.g.* at least about 400, such as at least about 1000, and preferably at least about 2300 bases thereof. In a specific embodiment of this aspect, the nucleic acid according to the invention comprises about the second half of the sequence of SEQ ID NO 1, preferably at least about 1200, *e.g.* 1284, and most preferably, the 1284 bases that encodes the protein denoted MSA1 which is discussed above. In another embodiment of this aspect of the invention, the nucleic acid is substantially identical with the sequence of SEQ ID NO 1. In a particularly advantageous embodiment, the present nucleic acid encodes a protein as disclosed by SEQ ID NO 3. In the present context, it is to

be understood that the invention also encompasses any one of the above defined sequences which is a degenerate or variant thereof.

Further, this aspect of the invention also covers any nucleic acid, which hybridizes specifically under stringent conditions to a nucleic acid as disclosed above as well as any gene product obtained thereby. Such a hybridizing nucleic acid may *e.g.* be DNA, a genomic DNA sequence comprising introns as well as exons, RNA, such as mRNA *etc.*

The preparation of the present nucleic acids has already been described above. Thus, more specifically, the sequences are cloned, or amplified by *in vitro* methods, such as the polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (SSR). A wide variety of cloning and *in vitro* amplification methodologies are known to persons of skill. Examples of these techniques and instructions sufficient to direct persons of skill through many cloning exercises are found in Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al.* (1989) *Molecular Cloning - A Laboratory Manual* (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, (Sambrook *et al.*); *Current protocols in Molecular Biology*, F.M. Ausubel *et al.*, Current Protocols, a joint venture between Greene Publishing Associates, Inc. And John Wiley & Sons, Inc., (1994 Supplement) (Ausubel); Cashion *et al.*, U.S. patent number 5,017,478; and Carr, European Patent No. 0,246,864. Examples of techniques sufficient to direct persons of skill through *in vitro* amplification methods are found in Berger, Sambrook, and Ausubel, as well as Mullis *et al.*, (1987) U.S. Patent No. 4,683,202; *PCR Protocols A Guide to Methods and Applications* (Innis *et al.* Eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (October 1, 1990) *C&EN* 36-47; *The Journal of NIH Research* (1991) 3: 81-94 (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173; Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874; Lomell *et al.* (1989) *J. Clin. Chem.*, 35: 1826; Landegren *et al.*,

(1988) *Science*, 241: 1077-1080; Van Brunt (1990) *Biotechnology*, 8: 291-294; Wu and Wallace, (1989) *Gene*, 4: 560; and Barringer *et al.* (1990) *Gene*, 89: 117.

In one preferred embodiment, the nucleic acids according to the invention are isolated by routine cloning methods. The DNA sequence provided in SEQ ID NO 1, or a subsequence thereof, can be used to provide probes that specifically hybridize to the encoding gene, in a genomic DNA sample, *e.g.* in a Southern blot, or to the mRNA, in a total RNA sample (*e.g.*, in a Northern blot). Once the target nucleic acid has been identified (*e.g.*, in a Northern or Southern blot), it is isolated according to standard methods known to those of skill in the art (*see, e.g.*, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Vols. 1-3, Cold Spring Harbor Laboratory; Berger and Kimmel (1987) *Methods in enzymology*, Vol. 152: *Guide to Molecular Cloning Techniques*, San Diego: Academic press, Inc.; or Ausubel *et al.* (1987) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York).

The nucleic acids according to the invention may be used in assays, such as in hybridization assays, as probes, in which case they are preferably labeled, in the production of proteins according to the invention *etc.*, as is evident from other sections of the present specification.

A further aspect of the invention is an expression vector, such as a plasmid or a virus, such as a phage, which comprises a nucleic acid as describe above. Such vectors are *e.g.* useful for expressing the proteins according to the invention to provide immunogens for antibody production. Thus, vectors encoding the proteins are also useful for transforming cells *in vitro* or *in vivo* to express the present protein or a functional subsequence thereof. In addition to the present encoding sequence, the vectors according to the invention also comprise suitable promoters, enhancers and further regulatory elements operably linked thereto (Queen *et al.* (1986) *Immunol. Rev.* 89:49). The vector may be a plasmid, virus, *etc.* Further, the vector may be an oligonucleotide, in which case the coding sequence may be

accompanied by the required tagging sequences for use in methods such as homologous recombination, as described in the literature.

Thus, an additional aspect of the present invention is a recombinant cell comprising a vector according to the invention. The culture of cells, including cell lines and cultured cells from tissue or blood samples, is well known in the art. (See *e.g.* Freshney, *Culture of Animal Cells, A Manual of Basic Technique*, 3rd ed., Wiley-Liss, New York, NY (1994).) Cells expressing the present nucleic acid may be used to monitor expression levels of the protein according to the invention in a wide variety of contexts. The cells according to the invention may be prokaryotic, such as bacteria, *e.g. E. coli*, eukaryotic, such as yeast, mammalian, such as canine, porcine or human, a protozoan *etc.*

In an additional aspect, the invention relates to a method for producing a protein according to the invention, or a functional subsequence thereof, which method comprises the steps of

- (a) providing a DNA encoding the desired protein or polypeptide;
- (b) introducing said DNA in a suitable expression vector or expression cassette;
- (c) transfer of said vector or cassette into a suitable cell;
- (d) culturing said cell to obtain the desired product; and optionally
- (e) purification of the protein or polypeptide.

The transfer of the vector may be performed by well-known methods, depending on the type of cellular host, such as calcium chloride transfection, which is commonly used with prokaryotic cells, or calcium phosphate treatment, electroporation, lipofection, microinjection *etc.* Once expressed, the products may be purified according to standard procedures in the art, such as HPLC purification, fraction column chromatography, gel electrophoresis and the like (see *e.g.* Scopes, *Protein Purification*, Springer-Verlag, NY, 1982).

In a further aspect, the present invention relates to an antibody raised against a protein according to the invention, or a functional part thereof. The antibody according to the invention will preferably comprise at least about 10, more preferably at least 20, 40 or 50 and most preferably at least 100 or 200, or even 400 amino acids. In a specific embodiment, the antibody binds to a protein comprised of essentially all of the sequence disclosed in SEQ ID NO 2.

More specifically, the present invention relates to antibodies including individual, allelic, strain, or species variants, and fragments thereof, both in their naturally occurring (full-length) forms and in recombinant forms. Additionally, antibodies are raised to these polypeptides in either their native configurations or in non-native configurations. Anti-idiotypic and chimeric or bispecific antibodies can also be generated.

In order to produce the antibodies specifically reactive with polypeptides according to the invention, a number of immunogens are used. Recombinant or synthetic polypeptides of 8-15, preferably 10, amino acids in length, or greater, selected from amino acid sub-sequences of SEQ ID NO 2 are the preferred polypeptide immunogen (antigen) for the production of monoclonal or polyclonal antibodies. In one class of preferred embodiments, an immunogenic peptide conjugate is also included as an immunogen. Naturally occurring polypeptides are also used, either in pure or impure form.

Recombinant polypeptides are expressed in eukaryotic or prokaryotic cells (as described above) and purified using standard techniques. The polypeptide, or a synthetic version thereof, is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies can be generated for subsequent use in immunoassays to measure the presence and quantity of the polypeptide.

Methods of producing polyclonal antibodies are known to those of skill in the art, see, e.g., Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY).

Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies are screened for binding to normal or modified polypeptides, or screened for agonistic or antagonistic activity, e.g., activity mediated through a *suppressor of fused* protein. Specific monoclonal and polyclonal antibodies will usually bind with a K_D of at least about 0.1 mM, more usually at least about 50 μ M, and most preferably at least about 1 μ M or better.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as dogs, swine, rodents, primates, humans, *etc.* Description of techniques for preparing such monoclonal antibodies are found in, e.g. Stites *et al.* (eds) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane, *supra*; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY; and Kohler and Milstein (1975) *Nature* 256: 495-497.

Other suitable techniques involve selection of libraries of recombinant antibodies in phage or similar vectors (see, e.g., Huse *et al.* (1989) *Science* 246: 1275-1281; and Ward, *et al.* (1989) *Nature* 341: 544-546; and Vaughan *et al.* (1996) *Nature Biotechnology*, 14: 309-314).

In a further aspect, the present invention relates to a protein as defined above, or a functional subsequence thereof, for use in a pharmaceutical preparation, preferably as a vaccine. Thus, the invention also relates to a vaccine preparation comprising a protein as defined above and a pharmaceutically acceptable carrier, such as an

aqueous carrier. A variety of aqueous carriers may be used, *e.g.* buffered saline. The solution is sterile and generally free of any undesirable matter. It may contain further auxiliary substances as required to approximate physiological conditions, such as pH adjusting agents and buffers, toxicity adjusting agents, *e.g.* sodium acetate, sodium chloride, potassium chloride, sodium lactate *etc.* It may be prepared for injection or any other suitable route of administration, such as oral administration. For a brief review of drug delivery, see Langer, Science 249:527-1533 (1990).

More specifically, the vaccine preparation according to the invention is capable of provoking an immune response in a subject, such as an animal or human, and thereby preventing said subject from being infected with *Sarcoptes scabiei*, whereby sarcoptes mange or scabies is avoided. Thus, the vaccine composition according to the invention may advantageously be used to immunize animals, *e.g.* dogs or swine.

As one example, dogs may be infected with scabies or sarcoptic mange from wild animals in the surroundings, such as red fox. Swine is another species that may be infected with scabies or sarcoptic mange. For example, when piglets are transported to novel locations during the breeding thereof, they may be contacted with other stock and the infection may then be spread. The vaccine composition may be a killed or inactivated, attenuated, recombinant or subunit vaccine, as appropriate, depending on the prevailing conditions.

Thus, the invention also relates to a method of preventing mite diseases, preferably conditions caused by mites, such as *Sarcoptes scabiei*, in a subject, such as a human, or animal, such as a canine or porcine, subject, which method comprises the administration of a pharmaceutical preparation according to the invention to said subject in an effective dose. In a specific embodiment, said disease is sarcoptes mange or scabies.

The invention also relates to any medicinal use of the proteins and polypeptides according to the invention. As regards the pharmaceutical uses of the present

polypeptides, it is to be understood, that in many cases, it may be more advantageous to use the peptidomimetics according to the invention than the original polypeptides. Such uses are also within the scope of the invention.

5 In one further aspect, the invention relates to various assays, wherein an antibody according to the invention is used to determine the presence and/or quantity of the polypeptide. The methods include analytical biochemical methods, such as electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography and the like as well as
10 various immunological methods, such as fluid or gel precipitating reactions, immunodiffusion, immunoelectrophoresis, radioimmunoassay (RIA), enzyme-linked immunosorbent assays (ELISAs), immunofluorescence assays *etc.* As used herein, an immunoassay is an assay wherein an antibody is used to specifically bind to the analyte. In the present context, such an assay format as disclosed in EP 291 194
15 (Unilever) is advantageously used, or variants of said format specifically adapted to the present purpose. For a review of general immunoassays, see *Methods in Cell Biology*, Vol. 37: *Antibodies in Cell Biology*, Asai, ed., Academic Press, Inc., New York (1993); and *Basic and Clinical Immunology*, 7th ed., Stites & Terr, eds., (1991).

20 Competitive assay formats are preferred in the present context, wherein the amount of analyte, preferably an unknown quantity of antibodies in a subject, in a sample is measured indirectly by measuring the amount of added analyte, displaced from a capture agent by the analyte present in the sample. Most preferred are the enzyme-
25 linked immunosorbent assay (ELISA) methods, in which an antibody typically is bound to an enzyme, such as peroxidase or phosphatase, which can produce colored reaction products from an appropriate buffer. Thus, it utilizes a tagged antigen molecule of known quantity to determine an unlabelled antigen of unknown quantity. Preferably, the protein according to the invention, or a suitable functional
30 fragment thereof, is used coupled to a conventional tag, such as His6. This assay is *e.g.* useful to diagnose *Sarcoptes scabiei* infection in dogs.

Thus, in an ELISA format according to the invention, antibodies against the present polypeptide are detected and/or quantified, preferably in a biological sample. The sample may be any sample of biological tissue or fluid, such as blood. The sample is pretreated as necessary by dilution in a suitable buffer solution or concentrated, if desired. Any number of standard aqueous buffer solutions may be used, such as Tris or the like, at physiological pH. Samples are incubated with an excess of the protein according to the invention as antigen. After rinsing to remove any unbound antibody, the amount of bound antibody is quantitated by adding a solution of enzyme-conjugated antibody that binds to constant domains of antibodies in the sample. Excess conjugated antibody is rinsed away and the activity of the bound enzyme is determined by adding the substrate to the reaction and measuring the formation of products. As the products of the reactions used in ELISA procedures are colored, the amount of product formed can readily be determined by the intensity of the colour that has developed using a spectrophotometer. The activity of the bound enzyme is proportional to the amount of antigen-binding antibody in the sample; therefore, the original concentration of such antibodies can be estimated from a series of control assays employing known concentrations of specific antibodies.

More specifically, if a protein as defined by SEQ ID NO. 2 coupled to His6 is used, it has been shown by the present inventor that an antigen concentration of about 30-60 ng/ml is advantageous. It is essential that protein contamination is prevented in order to avoid an enhanced background. Further, the coating ratio should be adapted as appropriate in order to improve the background. Presumably, the aftercoating is not critical.

Consequently, the invention also relates to a kit for the diagnosis of a subject infected with a mite, preferably *Sarcoptes scabiei*. More specifically, the kit according to the invention is adapted for performing the above disclosed ELISA method. Such a kit will include one or more reagents for determining the presence

or absence of antibodies in the subject raised against the polypeptides according to the invention. The antigen may be free or immobilized on a solid support, such as a test tube, a microtiter plate, a dipstick or the like. The kit may also comprise instructions for the use thereof. The kit may also contain means for the detection of labels, positive and negative controls, washing solutions *etc.* More specifically, a kit according to the present invention may advantageously comprise a conjugate, positive and negative control samples, a serum dilution buffer, a substrate, suitable washing solutions as well as an appropriate conjugate dilution buffer.

Detailed description of the drawings

Figure 1 shows the extent of the recombinant protein according to the invention in relation to the native protein.

Figure 2 illustrates the recombinant plasmid pPU17 used for the expression of the mite protein MSA1 according to the invention. Amp is the ampicillin resistance gene, lacI is the lac repressor gene, and malE is the maltose binding protein in fusion with the present MSA1.

Figure 3 is an overview of the cloning strategy of the 5' end of MSA.

Figure 4 shows the results of Western blot analysis of the mite recombinant protein MSA1 according to the invention. The protein was expressed in *E.coli*, purified and separated by SDS-PAGE before transfer to nitrocellulose. Lanes 1, 2 and 4 show the results after analysis with positive dog sera, and lanes 3, 5 and 6 shows the results after analysis with negative dog sera. Lane M is the molecular weight markers.

EXPERIMENTAL

Material and methods

Collection of parasites

Living mites of both sexes and different developmental stages of *S. scabiei* were isolated from the skin of wild red foxes as described in Bornstein & Zakrisson (1993) *Vet. Dermatol.* 4:107. Briefly, pieces of skin with the fur cut down to about 10 mm were left in Petri dishes at room temperature

under an electric light. Mites that migrated onto the underside of the lid were collected and stored at -70°C until further use.

Immunosera

5 *S. scabiei* proteins of a size of above 90 kD were size-selected on 10% gels by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and electroeluted as previously described, see Jacobs & Clad (1986) Anal. Biochem. 154:583. New Zealand White rabbits were immunized with the selected proteins and Freund's complete adjuvant.

RNA preparation and construction of cDNA library

10 A total of 160 mg of mites was washed with PBS and then homogenized in a glass grinder at 4°C . Total RNA was extracted with an RNAgents® kit from Promega and mRNA was isolated by oligo(dT) cellulose chromatography (Pharmacia Biotech). Double-stranded cDNA was
15 synthesized from about 5 μg *S. scabiei* mRNA (ZAP-cDNA, Stratagene), portions ligated into UNI-ZAP XR vector (Stratagene) and packaged into λ -phage (Gigapack Gold II, Stratagene).

Cloning and sequencing of MSA1

20 The amplified library was screened with sera from the rabbits previously immunized with gel purified antigens. Immuno reactive plaques were cloned and the phages eluted. From a positive λ -phage clone, a pBluescript SK-subclone was generated using the ExAssist™ phage system
25 (Stratagene). The resulting clone was designated pPU3 and the *S. scabiei* cDNA insert was mapped for positions of restriction endonuclease cleavage using enzymes from Pharmacia Biotech (Uppsala, Sweden) and New England Biolabs. For the sequencing of pPU3, nested deletions and cDNA sequence specific primers were used in conjunction with the T7 sequence
30 kit from Pharmacia Biotech.

Subcloning and expression of MSA1

Construction of high expression vector

The open reading frame of the recombinant plasmid pPU3 was amplified by PCR using *Pfu* DNA polymerase (Stratagene) and subcloned into the XmnI-XbaI sites of the vector pPU16. The pPU16 vector is a derivative of the pMal-c2 vector (New England Biolabs) that has a sequence of DNA encoding for 6 consecutive histidine residues and a stop codon between the PstI and HindIII site. The resulting plasmid pPU17 encodes for the recombinant mite protein with the maltose binding protein fused to its N-terminus and a His6-tag fused to its C-terminus. The junction between the maltose binding protein and the recombinant mite protein has a factor Xa-binding site (IEGR).

Expression as fused protein

The recombinant plasmid pPU17 was transformed into *E. coli* strain BL21 (DE3). The resulting transformant was inoculated into 10 ml minimal medium with casamino acids and heavy metals (MM/CA) (Pryor & Leiting, Protein Expr Purif 10:309) containing ampicillin (100 mg/ml) and grown overnight at 37°C with continuous shaking. Five ml of the overnight culture was diluted in 500 ml fresh MM/CA medium with ampicillin (100 mg/ml) and grown in a 2-liter Ehrlenmeyer flask shaking efficiently at 37°C until the OD value at 600 nm became 0.8. The culture was then cooled to 18°C, induced with 0.5 mM IPTG and then transferred to a shaker at 18°C for continued growth for 20 h. After the expression the cells were collected by centrifugation at 4000xg for 20 min and resuspended in 25 ml 20 mM phosphate buffer (pH 7.6) with 0.5 M NaCl and 30 mM imidazole, supplemented with the Complete™ protease inhibitor (Roche) and frozen overnight at -70°C. The frozen cell suspension was thawed in cold water, placed in an ice-water bath and sonicated. The suspension was clarified by centrifugation at 9000xg for 30 min and the resulting supernatant was used for affinity purification.

Minimal medium with casamino acids and heavy metal (MM/CA) is: 5 g/l glucose, 1 mg/l (+) biotin, 2 mg/l thiamine, 1 g/l $(\text{NH}_4)_2\text{SO}_4$, 4 ml of a 250xheavy metal stock solution, 50 ml sterile 10% casamino acids, and 200 ml sterile 5xphosphate buffer. One liter of 250xheavy metal stock solution is 500 mg $\text{MoNa}_2\text{O}_4 \cdot 2\text{H}_2\text{O}$, 250 mg CoCl_2 , 175 mg $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1 g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 8.75 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.25 g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 1.25 g $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$, 2.5 g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, and 1.0 g H_3BO_3 in 1 M HCl. One liter of 5xphosphate buffer contains 53 g K_2HPO_4 and 24.7 g KH_2PO_4 .

Purification of recombinant mite antigen fusion protein

The supernatant was loaded onto a 1 ml HiTrap® chelating column (Amersham Pharmacia Biotech) loaded with N^{2+} at a flow rate of 1 ml/min. After washing with 10 ml of 20 mM phosphate buffer (pH 7.6) with 0.5 M NaCl and 30 mM imidazole, the captured recombinant protein was eluted with 3 ml 20 mM phosphate buffer (pH 7.6) with 0.5 M NaCl and 500 mM imidazole. After buffer exchange to 20mM Tris-HCl (pH 8.0) with 150 mM NaCl and 1mM EDTA, the maltose binding protein was cleaved off with factor Xa at room temperature for 20 h after which the recombinant mite protein was purified on a HiTrap® Q column. As preserver Micro-O-Protect (Roche) was added to the purified recombinant protein in concentraion of 0.1 %.

Cloning of part of the 5' cDNA end of MSA1

A PCR strategy was used in order to clone regions upstream of the cDNA insert in pPU3. In the first PCR the primer KBE 5 (5'CAC TAT CGG AGA ACG TAA CTT CGG 3'), complementary to the anti-sense strand of the insert in pPU3, was desigend and used together with a T3 primer, complementary to the vector used to construct the cDNA library. As a template the *S. scabiei* cDNA-libray was used. The resulting fragment was cloned into the *Sma*I-site of pUC18 and sequenced as above. This new fragment was then used for the design of a second primer KBE 8 (5'CCT GGC ATT CTA CTT GAG ATG TA 3') for the amplification an additional 5'end cDNA fragment. The second 5'end fragment was cloned and sequenced as

above. A continuous cDNA which included the original MSA1 cDNA and both of the 5' end fragments was generated by using the Titan™ One Tube RT-PCR system (manufactured by Roche). For the reverse transcriptase step the reverse primer MSA1Xba (5'CGC **TCT AGA** CTC AAC AAT GAA TGT CTG CAA 3') was used. In the PCR, the reverse primer was used in combination with the forward primer LDL 2 (5'CGG **GAT CCG** AAT ATT TCG TCT CGA AAC CG 3'). The resulting fragment was cloned into the *Bam*HI-*Xba*I sites of pPU16 utilizing the recognition sites introduced during the PCR (shown in boldface). A graphic overview of the cloning strategy is shown in Fig.

Enzyme-linked immunosorbent assay

The recombinant mite protein MSA1 was diluted in a 0.1 M carbonate buffer, pH 9.6, to a concentration of 62.5 ng/ml and coated overnight at 4°C onto microtitre plates (Polysorp, manufactured by Nunc) in volumes of 100 µl per well. The plates were washed once with phosphate buffered saline with 0.05% Tween20 (PBS-T) followed by the addition of serum samples at a 1:100 dilution in PBS-T. After a 1h incubation at 37°C the plates were washed three times with PBS-T and a monoclonal mouse anti-dog IgG antibody, dilution 1:1000, and rabbit anti-mouse Ig conjugated to HRP (manufactured by Dako), dilution 1:1000, in PBS-T plus 1% normal rabbit serum were added to the wells and incubated for 1h at 37°C. After washing three times with PBS-T, 100µl of TMB (manufactured by Sigma) were added to each well. The substrate incubation was stopped with 50 µl of H₂SO₄ and the amount of end product was analyzed at 450 nm in Dynatech MR5000 spectrophotometer (manufactured by Dynatech).

Affinity purification of antibodies

Purified recombinant mite protein or recombinant phages in *E. coli* expressing the recombinant mite protein were transferred to a nitro cellulose filter. After blocking the filter with 1% BSA (Stratagene) in 20 mM Tris-HCl, 0.5 M NaCl, 0.05% Tween20, pH 7.5 (TTBS), anti-*S. scabiei* rabbit serum was added to the filter and incubated for 1h at room temperature. After extensive washing with TTBS, the

bound antibodies were eluted and used in Western blot analysis on *S. scabiei* antigens(Beall & Mitchell, J. Immunol. Method. (1986) 86:217). The affinity purified antibodies were also used to localize the protein in *S. scabiei* mites.

5 Western blot analysis

Purified recombinant mite proteins were separated by SDS-PAGE in 12% mini gels in Mini PROTEAN II cell (Bio-Rad Laboratories). Blotting of proteins to nitrocellulose was done by using an electrophoretic transfer cell (Bio-Rad Laboratories). After completion of transfer, the membrane bound
10 proteins were checked by briefly incubating the membrane in a 0.2% Ponceau-S solution (Sigma) followed by rinsing in distilled water. The blots were blocked with phosphate buffer saline (PBS) containing non-fat dry milk, incubated with sera from experimentally dogs (Bornstein & Zakrisson (1993) *Vet. Dermatol.* 4:107), then incubated with a mouse anti-dog IgG
15 monoclonal antibody and a rabbit anti-mouse IgG. After washing, bound antibodies were visualized by chemiluminescence detection and exposure to film using the ECL-system (Amersham Pharmacia Biotech).

SEQUENCE LISTING

<110> NATIONAL VETERINARY INSTITUTE

5 <120> MITE PROTEIN

<130> 53631

<140>

10 <141>

<160> 3

<170> PatentIn Ver. 2.1

15

<210> 1

<211> 2313

<212> DNA

<213> Sarcoptes scabiei

20

<220>

<221> CDS

<222> (1)..(2310)

25

<220>

<221> CDS

<222> (1030)..(1278)

<400> 1

30

gaa tat ttc gtc tcg aaa ccg aac aaa gat gtc aat gcg atc gaa ttg 48
 Glu Tyr-Phe Val Ser Lys Pro Asn Lys Asp Val Asn Ala Ile Glu Leu
 1 5 10 15

35

gaa ttc aaa cac gag agt gat gat tct aaa aag aat cga aaa tat acc 96
 Glu Phe Lys His Glu Ser Asp Asp Ser Lys Lys Asn Arg Lys Tyr Thr
 20 25 30

40

gcc gaa ttg aga gaa gtt ggt ccg gct acc ccg aaa acg gca aaa ctt 144
 Ala Glu Leu Arg Glu Val Gly Pro Ala Thr Pro Lys Thr Ala Lys Leu
 35 40 45

45

gag atc gat gtt gcc aaa gga gag gaa tac aaa gtg aca atg aaa tca 192
 Glu Ile Asp Val Ala Lys Gly Glu Glu Tyr Lys Val Thr Met Lys Ser
 50 55 60

ccg aac aat gaa ttc cat act gaa ttc act ttt gct gcc gat aag aat 240
 Pro Asn Asn Glu Phe His Thr Glu Phe Thr Phe Ala Ala Asp Lys Asn
 65 70 75 80

	cat ttg aaa atg aaa gcc gat ttt ccg gat aga ttc cgt gcc gat gtc	288
	His Leu Lys Met Lys Ala Asp Phe Pro Asp Arg Phe Arg Ala Asp Val	
	85 90 95	
5	acc ggt aca ttc gag cat gat aaa gag act ggt gta cgt aaa aac aaa	336
	Thr Gly Thr Phe Glu His Asp Lys Glu Thr Gly Val Arg Lys Asn Lys	
	100 105 110	
10	ttg aat gtc gaa tac aaa ttg ggt agc gat gat aaa gct cat aca att	384
	Leu Asn Val Glu Tyr Lys Leu Gly Ser Asp Asp Lys Ala His Thr Ile	
	115 120 125	
15	gaa tat gag aac gag atg gct ttc aaa ttg aaa cga tca tcg aaa gag	432
	Glu Tyr Glu Asn Glu Met Ala Phe Lys Leu Lys Arg Ser Ser Lys Glu	
	130 135 140	
20	aag aat aca aat ttg atg tat aaa tcc aaa tac atc tca agt aga atg	480
	Lys Asn Thr Asn Leu Met Tyr Lys Ser Lys Tyr Ile Ser Ser Arg Met	
	145 150 155 160	
25	cca ggt cta aat cat aaa aca gcg ctg gaa ttc aaa tat cga cca ttc	528
	Pro Gly Leu Asn His Lys Thr Ala Leu Glu Phe Lys Tyr Arg Pro Phe	
	165 170 175	
30	aaa acg aat gat ttg tat cta gaa ttg gaa ttc ggt aat gat tta caa	576
	Lys Thr Asn Asp Leu Tyr Leu Glu Leu Glu Phe Gly Asn Asp Leu Gln	
	180 185 190	
35	cac aaa tat caa ttg caa cga aaa acc gat atg gaa gtg gaa gag atg	624
	His Lys Tyr Gln Leu Gln Arg Lys Thr Asp Met Glu Val Glu Glu Met	
	195 200 205	
40	cga cca ttt aaa ttg aaa gga aat agt gat atc aaa ttg gtt gcg acc	672
	Arg Pro Phe Lys Leu Lys Gly Asn Ser Asp Ile Lys Leu Val Ala Thr	
	210 215 220	
45	gat ttc gat gtc gat tac gat ctc aaa tcg gat ttc aaa tat gaa agc	720
	Asp Phe Asp Val Asp Tyr Asp Leu Lys Ser Asp Phe Lys Tyr Glu Ser	
	225 230 235 240	
50	aat aaa ggt acg ccg atg gaa ttg caa tac aat ctc aaa ggt aaa gat	768
	Asn Lys Gly Thr Pro Met Glu Leu Gln Tyr Asn Leu Lys Gly Lys Asp	
	245 250 255	
55	cga tca aaa cgg gcg gcg gaa aag aat caa gag gaa atc gaa gga aag	816
	Arg Ser Lys Arg Ala Ala Glu Lys Asn Gln Glu Glu Ile Glu Gly Lys	
	260 265 270	

	att gat tac aaa aat aat gga tct ccg atc gat tca aaa atg aat gcc Ile Asp Tyr Lys Asn Asn Gly Ser Pro Ile Asp Ser Lys Met Asn Ala 275 280 285	864
5	aat ctc caa gct tgg ggt aat caa tat gct tac gaa agt gaa ttg aaa Asn Leu Gln Ala Trp Gly Asn Gln Tyr Ala Tyr Glu Ser Glu Leu Lys 290 295 300	912
10	caa gtt gaa cca caa cga tat gag gga aag att acg atg tcg aaa aat Gln Val Glu Pro Gln Arg Tyr Glu Gly Lys Ile Thr Met Ser Lys Asn 305 310 315 320	960
15	gat aag aag atc ttc atc act cac aaa gat gaa atg gct aaa cca acc Asp Lys Lys Ile Phe Ile Thr His Lys Asp Glu Met Ala Lys Pro Thr 325 330 335	1008
20	gat aca ttc cat ctg aaa tcc gaa gcg gaa gtt acg ttc tcc gat agt Asp Thr Phe His Leu Lys Ser Glu Ala Glu Val Thr Phe Ser Asp Ser 340 345 350	1056
	gaa gat aag aaa aat tat ttc gtc gaa ctt aaa aaa gat aaa gat tta Glu Asp Lys Lys Asn Tyr Phe Val Glu Leu Lys Lys Asp Lys Asp Leu 355 360 365	1104
25	tat tcg atg aaa tcg aat gtg aaa cga aac aat gag att ttc tat gag Tyr Ser Met Lys Ser Asn Val Lys Arg Asn Asn Glu Ile Phe Tyr Glu 370 375 380	1152
30	aac aat atg gat ttg gag aag aac ggt aaa atg aat tgg tat tac aaa Asn Asn Met Asp Leu Glu Lys Asn Gly Lys Met Asn Trp Tyr Tyr Lys 385 390 395 400	1200
35	cga aac gat cga aca tgg aat atg gat ctc gat aat gca ttc aat cca Arg Asn Asp Arg Thr Trp Asn Met Asp Leu Asp Asn Ala Phe Asn Pro 405 410 415	1248
40	aga gat ggt aca atg aaa ctt caa gtg aaa gat cgt atc tat gat atc Arg Asp Gly Thr Met Lys Leu Gln Val Lys Asp Arg Ile Tyr Asp Ile 420 425 430	1296
	aaa ttg aaa cga gaa ccg ttc cga tac ggt gat cta cat atc gaa gga Lys Leu Lys Arg Glu Pro Phe Arg Tyr Gly Asp Leu His Ile Glu Gly 435 440 445	1344
45	aat gag aat cct ttg atc aaa aag ggt gat tta cat atg tct ctc gtc Asn Glu Asn Pro Leu Ile Lys Lys Gly Asp Leu His Met Ser Leu Val 450 455 460	1392

	gat ccg ctt act ttg aat gtt ttg acc aag aat gat gga atc gtc gat	1440
	Asp Pro Leu Thr Leu Asn Val Leu Thr Lys Asn Asp Gly Ile Val Asp	
	465 470 475 480	
5	atg aca ttg gat ttg gtc tct ccc aac acc aaa aaa gca gcg cta aaa	1488
	Met Thr Leu Asp Leu Val Ser Pro Asn Thr Lys Lys Ala Ala Leu Lys	
	485 490 495	
10	atc aat tcg aaa aaa tac gat ctt gat cat gat ggt gag att acc gtt	1536
	Ile Asn Ser Lys Lys Tyr Asp Leu Asp His Asp Gly Glu Ile Thr Val	
	500 505 510	
15	tcg atc ttt aat cct cga atg act tgg aaa cat cac act aga aaa ggt	1584
	Ser Ile Phe Asn Pro Arg Met Thr Trp Lys His His Thr Arg Lys Gly	
	515 520 525	
20	gat atg gaa ttg aat att gat gct gat atc act cga aaa ggt tca ttg	1632
	Asp Met Glu Leu Asn Ile Asp Ala Asp Ile Thr Arg Lys Gly Ser Leu	
	530 535 540	
25	atc acc tat tct cgt aaa gag cca gat gat tcg aca aaa gtt cga tat	1680
	Ile Thr Tyr Ser Arg Lys Glu Pro Asp Asp Ser Thr Lys Val Arg Tyr	
	545 550 555 560	
30	tca aga caa gga aat caa gtt tcg atg gaa gtc gat tct aaa ttg atc	1728
	Ser Arg Gln Gly Asn Gln Val Ser Met Glu Val Asp Ser Lys Leu Ile	
	565 570 575	
35	gaa ggc cat gcg aac gga act ttg acc gat ggc aaa att cat gtc aaa	1776
	Glu Gly His Ala Asn Gly Thr Leu Thr Asp Gly Lys Ile His Val Lys	
	580 585 590	
40	ggt cga gag agt gat ttc gaa atc gaa agc acc tat aaa gtt gaa gat	1824
	Gly Arg Glu Ser Asp Phe Glu Ile Glu Ser Thr Tyr Lys Val Glu Asp	
	595 600 605	
45	ggt aag ctt atg att gag cca acc aaa act cag aat gga aaa tta gaa	1872
	Gly Lys Leu Met Ile Glu Pro Thr Lys Thr Gln Asn Gly Lys Leu Glu	
	610 615 620	
50	ggt ctt ctt tcg aga aaa gta cca tca cat ctt gtt ctt gaa aca cca	1920
	Gly Leu Leu Ser Arg Lys Val Pro Ser His Leu Val Leu Glu Thr Pro	
	625 630 635 640	
55	aga gtg aaa atg aac atg aaa tat gat aga ttt gct ccg gtg aag ata	1968
	Arg Val Lys Met Asn Met Lys Tyr Asp Arg Phe Ala Pro Val Lys Ile	
	645 650 655	

	ttg aaa tta gat tac gat ggt ttg aat tat gag aaa cat atc gat gct	2016
	Leu Lys Leu Asp Tyr Asp Gly Leu Asn Tyr Glu Lys His Ile Asp Ala	
	660 665 670	
5	gaa tac gag cca tca aat cat tac aaa tac ttt acc gat ggt aaa tcg	2064
	Glu Tyr Glu Pro Ser Asn His Tyr Lys Tyr Phe Thr Asp Gly Lys Ser	
	675 680 685	
10	aag aga agt ggc aaa ggt tat tcg atc aaa atc gat gga aaa cca aag	2112
	Lys Arg Ser Gly Lys Gly Tyr Ser Ile Lys Ile Asp Gly Lys Pro Lys	
	690 695 700	
15	aaa gca ttg aaa gtt gat gtc gat atg ccg gat ttt aaa ttc aat gtg	2160
	Lys Ala Leu Lys Val Asp Val Asp Met Pro Asp Phe Lys Phe Asn Val	
	705 710 715 720	
20	aac aaa ccg gaa gat agt aac aaa gct caa ttt agt tat aca ttc aat	2208
	Asn Lys Pro Glu Asp Ser Asn Lys Ala Gln Phe Ser Tyr Thr Phe Asn	
	725 730 735	
25	gat tat acc gaa acg gaa gag tat gaa ttc gat cca cat cgt gca tat	2256
	Asp Tyr Thr Glu Thr Glu Glu Tyr Glu Phe Asp Pro His Arg Ala Tyr	
	740 745 750	
30	atc ttg aat tgg gcc aga gct atc aga caa tat ttg cag aca ttc att	2304
	Ile Leu Asn Trp Ala Arg Ala Ile Arg Gln Tyr Leu Gln Thr Phe Ile	
	755 760 765	
35	gtt gag gatcgtatct atgatatcaa attgaaacga gaaccgttcc gatacgggtga	2360
	Val Glu	
	770	
40	tctacatatc gaaggaaatg agaatccttt gatcaaaaag ggtgatttac atatgtctct	2420
	cgtcgatccg ctactttga atgttttgac caagaatgat ggaatcgtcg atatgacatt	2480
	ggatttggtc tctccaaca ccaaaaaagc agcgctaaaa atcaattcga aaaaatacga	2540
	tcttgatcat gatggtgaga ttaccgtttc gatctttaat cctcgaatga cttggaaaca	2600
	tcacactaga aaaggtgata tggaattgaa tattgatgct gatatactc gaaaagggtc	2660
	attgatcacc tattctcgta aagagccaga tgattcgaca aaagttcgat attcaagaca	2720
45	aggaaatcaa gtttcgatgg aagtcgattc taaattgatc gaaggccatg cgaacggaac	2780
	tttgaccgat ggcaaaattc atgtcaaagg tcgagagagt gatttcgaaa tcgaaagcac	2840
	ctataaagtt gaagatggta agcttatgat tgagccaacc aaaactcaga atggaaaatt	2900

agaaggtctt ctttcgagaa aagtaccatc acatcttggt cttgaaacac caagagtgaa 2960
 aatgaacatg aaatatgata gatttgctcc ggtgaagata ttgaaattag attacgatgg 3020
 5 tttgaattat gagaaacata tcgatgctga atacgagcca tcaaatcatt acaaatactt 3080
 taccgatggg aaatcgaaga gaagtggcaa aggttattcg atcaaaatcg atggaaaacc 3140
 10 aaagaaagca ttgaaagttg atgtcgatat gccggatttt aaattcaatg tgaacaaacc 3200
 ggaagatagt aacaaagctc aatttagtta tacattcaat gattataccg aaacggaaga 3260
 15 gtatgaattc gatccacatc gtgcatatat ctigaattgg gccagagcta tcagacaata 3320
 tttgcagaca ttcattgttg agtag 3345

 20 <210> 2
 <211> 770
 <212> PRT
 <213> *Sarcoptes scabiei*

 25 <400> 2
 Glu Tyr Phe Val Ser Lys Pro Asn Lys Asp Val Asn Ala Ile Glu Leu
 1 5 10 15
 Glu Phe Lys His Glu Ser Asp Asp Ser Lys Lys Asn Arg Lys Tyr Thr
 30 20 25 30
 Ala Glu Leu Arg Glu Val Gly Pro Ala Thr Pro Lys Thr Ala Lys Leu
 35 35 40 45
 Glu Ile Asp Val Ala Lys Gly Glu Glu Tyr Lys Val Thr Met Lys Ser
 50 55 60
 Pro Asn Asn Glu Phe His Thr Glu Phe Thr Phe Ala Ala Asp Lys Asn
 40 65 70 75 80
 His Leu Lys Met Lys Ala Asp Phe Pro Asp Arg Phe Arg Ala Asp Val
 85 90 95
 Thr Gly Thr Phe Glu His Asp Lys Glu Thr Gly Val Arg Lys Asn Lys
 45 100 105 110
 Leu Asn Val Glu Tyr Lys Leu Gly Ser Asp Asp Lys Ala His Thr Ile
 115 120 125

	Glu Tyr Glu Asn Glu Met Ala Phe Lys Leu Lys Arg Ser Ser Lys Glu	
	130	135 140
5	Lys Asn Thr Asn Leu Met Tyr Lys Ser Lys Tyr Ile Ser Ser Arg Met	
	145	150 155 160
	Pro Gly Leu Asn His Lys Thr Ala Leu Glu Phe Lys Tyr Arg Pro Phe	
		165 170 175
10	Lys Thr Asn Asp Leu Tyr Leu Glu Leu Glu Phe Gly Asn Asp Leu Gln	
		180 185 190
	His Lys Tyr Gln Leu Gln Arg Lys Thr Asp Met Glu Val Glu Glu Met	
		195 200 205
15	Arg Pro Phe Lys Leu Lys Gly Asn Ser Asp Ile Lys Leu Val Ala Thr	
		210 215 220
20	Asp Phe Asp Val Asp Tyr Asp Leu Lys Ser Asp Phe Lys Tyr Glu Ser	
		225 230 235 240
	Asn Lys Gly Thr Pro Met Glu Leu Gln Tyr Asn Leu Lys Gly Lys Asp	
		245 250 255
25	Arg Ser Lys Arg Ala Ala Glu Lys Asn Gln Glu Glu Ile Glu Gly Lys	
		260 265 270
	Ile Asp Tyr Lys Asn Asn Gly Ser Pro Ile Asp Ser Lys Met Asn Ala	
		275 280 285
30	Asn Leu Gln Ala Trp Gly Asn Gln Tyr Ala Tyr Glu Ser Glu Leu Lys	
		290 295 300
	Gln Val Glu Pro Gln Arg Tyr Glu Gly Lys Ile Thr Met Ser Lys Asn	
35		305 310 315 320
	Asp Lys Lys Ile Phe Ile Thr His Lys Asp Glu Met Ala Lys Pro Thr	
		325 330 335
40	Asp Thr Phe His Leu Lys Ser Glu Ala Glu Val Thr Phe Ser Asp Ser	
		340 345 350
	Glu Asp Lys Lys Asn Tyr Phe Val Glu Leu Lys Lys Asp Lys Asp Leu	
		355 360 365
45	Tyr Ser Met Lys Ser Asn Val Lys Arg Asn Asn Glu Ile Phe Tyr Glu	
		370 375 380

	Asn Asn Met Asp Leu Glu Lys Asn Gly Lys Met Asn Trp Tyr Tyr Lys	
	385	390 395 400
5	Arg Asn Asp Arg Thr Trp Asn Met Asp Leu Asp Asn Ala Phe Asn Pro	
	405	410 415
	Arg Asp Gly Thr Met Lys Leu Gln Val Lys Asp Arg Ile Tyr Asp Ile	
	420	425 430
10	Lys Leu Lys Arg Glu Pro Phe Arg Tyr Gly Asp Leu His Ile Glu Gly	
	435	440 445
	Asn Glu Asn Pro Leu Ile Lys Lys Gly Asp Leu His Met Ser Leu Val	
15	450	455 460
	Asp Pro Leu Thr Leu Asn Val Leu Thr Lys Asn Asp Gly Ile Val Asp	
	465	470 475 480
20	Met Thr Leu Asp Leu Val Ser Pro Asn Thr Lys Lys Ala Ala Leu Lys	
	485	490 495
	Ile Asn Ser Lys Lys Tyr Asp Leu Asp His Asp Gly Glu Ile Thr Val	
	500	505 510
25	Ser Ile Phe Asn Pro Arg Met Thr Trp Lys His His Thr Arg Lys Gly	
	515	520 525
	Asp Met Glu Leu Asn Ile Asp Ala Asp Ile Thr Arg Lys Gly Ser Leu	
30	530	535 540
	Ile Thr Tyr Ser Arg Lys Glu Pro Asp Asp Ser Thr Lys Val Arg Tyr	
	545	550 555 560
35	Ser Arg Gln Gly Asn Gln Val Ser Met Glu Val Asp Ser Lys Leu Ile	
	565	570 575
	Glu Gly His Ala Asn Gly Thr Leu Thr Asp Gly Lys Ile His Val Lys	
	580	585 590
40	Gly Arg Glu Ser Asp Phe Glu Ile Glu Ser Thr Tyr Lys Val Glu Asp	
	595	600 605
	Gly Lys Leu Met Ile Glu Pro Thr Lys Thr Gln Asn Gly Lys Leu Glu	
45	610	615 620
	Gly Leu Leu Ser Arg Lys Val Pro Ser His Leu Val Leu Glu Thr Pro	
	625	630 635 640

	Arg Val Lys Met Asn Met Lys Tyr Asp Arg Phe Ala Pro Val Lys Ile	645	650	655
5	Leu Lys Leu Asp Tyr Asp Gly Leu Asn Tyr Glu Lys His Ile Asp Ala	660	665	670
	Glu Tyr Glu Pro Ser Asn His Tyr Lys Tyr Phe Thr Asp Gly Lys Ser	675	680	685
10	Lys Arg Ser Gly Lys Gly Tyr Ser Ile Lys Ile Asp Gly Lys Pro Lys	690	695	700
	Lys Ala Leu Lys Val Asp Val Asp Met Pro Asp Phe Lys Phe Asn Val	705	710	715
15	Asn Lys Pro Glu Asp Ser Asn Lys Ala Gln Phe Ser Tyr Thr Phe Asn	725	730	735
	Asp Tyr Thr Glu Thr Glu Glu Tyr Glu Phe Asp Pro His Arg Ala Tyr	740	745	750
20	Ile Leu Asn Trp Ala Arg Ala Ile Arg Gln Tyr Leu Gln Thr Phe Ile	755	760	765
25	Val Glu	770		
	<210> 3			
	<211> 83			
30	<212> PRT			
	<213> <i>Sarcoptes scabiei</i>			
	<400> 3			
35	Glu Ala Glu Val Thr Phe Ser Asp Ser Glu Asp Lys Lys Asn Tyr Phe	1	5	10
				15
	Val Glu Leu Lys Lys Asp Lys Asp Leu Tyr Ser Met Lys Ser Asn Val	20	25	30
40	Lys Arg Asn Asn Glu Ile Phe Tyr Glu Asn Asn Met Asp Leu Glu Lys	35	40	45
	Asn Gly Lys Met Asn Trp Tyr Tyr Lys Arg Asn Asp Arg Thr Trp Asn	50	55	60
45	Met Asp Leu Asp Asn Ala Phe Asn Pro Arg Asp Gly Thr Met Lys Leu	65	70	75
				80
	Gln Val Lys			